## Yukon River Fall Chum Salmon Integrated Population Model

### Authors:

Genoa H. Sullaway1\*, Curry J. Cunningham1, Lauren Rogers2, Sabrina Garcia3

Affiliations**:** 1 University of Alaska Fairbanks, College of Fisheries and Ocean Sciences, 17101 Point Lena Loop Rd., Juneau, AK 99801, 2National Oceanic and Atmospheric Administration, National Marine Fisheries Service, Alaska Fisheries Science Center, 7600 Sand Point Way NE, Seattle, WA 98115, USA 3 Alaska Department of Fish and Game

\* Corresponding Author

Contact Information: [gsullaway@alaska.edu](mailto:gsullaway@alaska.edu)

Running Page Head:

Keywords:

### 2. Methods

#### 2.1 Population Dynamics Model

An integrated population model (IPM) was used to estimate the influence of environmental and ecological covariates on the survival of Yukon River Chum salmon at multiple stages throughout their lifecycle (Figure 1, Figure 2). IPMs are a class of models, also called lifecycle models, that are well-suited to address the challenge of incorporating environmental predictors into population models (Schaub & Abadi 2011), as demonstrated by successful applications in mammal, bird and fish conservation (Besbeas et al. 2002, Cunningham et al. 2018, Regehr et al. 2018, Crozier et al. 2021, DeFilippo et al. 2021). The flexible IPM framework allowed incorporation of multiple data sources when estimating ecosystem covariate impacts on Fall Chum salmon survival. The IPM tracked cohorts of fall Chum salmon by brood year, *t*, life stage, *s* and age, *a*. The model includes four stages for Yukon river fall chum: 1) “eggs”, which tracks the amount of eggs produced by spawners, 2) “juvenile” which tracks individuals from eggs to the end of their first summer in the marine environment, 3) “total returns”, which tracks individuals from their first winter to when they return to the Yukon river mouth and are vulnerable to terminal harvest, , 4) “spawners”, which are the fish that return back to the spawning grounds and is the difference between total returns and total terminal harvest (Figure 2).

The number of Chum salmon surviving from an egg to the end of their first ocean summer, Nt,s=j depended upon the number of eggs spawned in brood year, *t*, Nt,s=e and the survival rate from eggs to ocean juveniles, t,s=j.

Nt,s=j= t,s=j\* Nt,s=e Eq. 4.1

The survival rate, t,s=j, was calculated using a Beverton-Holt transition function (Moussalli & Hilborn 1986).

Eq. 4.2

where the productivity parameter represents time-varying maximum survival rate in the absence of density-dependent compensation, and represents the carrying capacity, or the maximum number of individuals that could survive to the end of each life stage. The productivity parameter was estimated conditional on environmental covariates (Table 1) using an inverse logit function of basal productivity, , which represented the mean survival rate (in logit space) at low density.

Eq. 4.3

Here, a matrix of mean-scaled covariate values *c*, were multiplied by an associated coefficient which described the influence of each covariate, *c*, on stage specific survival rates.

Upon surviving their first summer at sea, Chum salmon migrate to the Eastern Aleutian Islands and Western Gulf of Alaska, where they spend up to five years at sea before returning to the Yukon River (Farley et al. 2024). The first winter in the GOA is hypothesized to be a critical life stage step. Thus, survival was estimated during the first winter at sea, t,s=m.  Survival during the first winter at sea was estimated using the Beverton-Holt transition function described above (Eq. 4.2) and the productivity parameter, which informs the maximum survival rate, was estimated conditional on environmental covariates described in Table 1 for the marine stage. Survival from the first winter at sea, t,s=m, was multiplied by the number of juvenile fish, Nt,s=j, to yield the number of fish that survive their first winter at sea, .

= t,s=m \* Nt,s=j Eq. 4.4

The number of fish returning to the Yukon River are based on calendar year, which is indexed by . The number of returning fish, Ny,s=r,a, depended on age structured natural mortality rates, and the proportion of fish that return to spawn from each brood year, .

= Eq. 4.5

We assumed a fixed natural mortality for ages 4-6, , where the annual mortality rate was 0.06. This represents the assumption that older fish had a higher marine mortality than younger fish (Beamish 2018), but that overall ocean mortality after the first winter at sea was low. The maturity schedule for Chum salmon was assumed to vary over time randomly, relative to an average maturity schedule. The proportion of fish returning to the Yukon River at each ocean age in from each brood year , was estimated as a Dirichlet hyper-distribution arising from a mean age at maturity probability vector, . Annual deviations from the mean age at maturity were determined by an inverse dispersion parameter, *D*.

Eq. 4.6

Returning fish, Ny,s=r,a, were subject to terminal harvest determined by annual fishing mortality in each calendar year , , and age-specific selectivity, . The resulting catch at age in calendar year is:

Eq. 4.7

To allow ample flexibility in annual fishing mortality rates by calendar year , , we estimated mean fishing mortality and process deviations around the mean, .

Eq. 4.8

Returning fish that were not captured in terminal fisheries were assumed to reach the spawning grounds and reproduce.

Eq. 4.9

The number of eggs produced by each spawner, *,* was assumed to follow a Ricker function which includes parameters capturing the log of the maximum recruitment per spawner without density dependence, , and the strength of density dependence, (Ricker 1954, Hilborn 1985). The proportion of females, , was fixed at 50% (Gilk et al.2009).

Eq. 4.10

The lifecycle begins again by summing the number of eggs produced by each spawner across age classes, yielding the eggs produced in each brood year, .

Eq. 4.11

#### 2.2 Model Estimation

We fit the IPM to multiple datasets using Bayesian inference and implemented the model in STAN (Carpenter et al. 2017) using the rstan package in R (Stan Development Team 2024). The model was fit using 4 chains, 8000 warm up iterations and 24,000 total iterations. We used an adapt-delta of 0.99 to force the model to take smaller steps when searching the parameter space. We diagnosed chain convergence using the Gelman-Rubin statistic (Brooks & Gelman 1998) and visually inspected coherence in the distribution of posterior samples from each chain using trace plots (Figure S2). We used the *priorsense* package in R to evaluate how sensitive the posterior distribution was to prior and likelihood perturbations (Kallioinen et al. 2023).

#### 2.3 Population Data

Marine juvenile abundance estimated in the IPM (Nt,s=j) was fit to a juvenile fall Chum salmon index arising from data collected in the Northern Bering Sea (NBS) survey. The NBS survey is a collaborative survey run by ADFG, NOAA Alaska Fishery Science Center to better understand the Northern Bering Sea Ecosystem (Murphy et al. 2021). The survey has collected a wide array of ecosystem information, including juvenile salmon abundance from surface trawls that are conducted at multiple stations across the NBS. The survey is conducted annually in Summer/Fall (typically between August and September), which is the termination of juvenile Chum salmon first summer at sea before they migrate to the Gulf of Alaska/Eastern Aleutians.

To account for spatial and temporal survey imbalances, Chum salmon CPUE data were used to estimate a juvenile Chum salmon index (**Garcia et al in prep?**). A Vector Autoregressive Spatio-Temporal modeling approach was used to create an independent index of juvenile chum salmon, and methods are detailed by Cunningham et al (Thorson 2019). The NBS survey includes multiple stocks of juvenile chum salmon that spend their first summer in the NBS. Chum salmon caught in this survey are allocated to genetic reporting groups using a spatial mixed stock analysis (MSA) (**CITE**, **Lee et al..?).** The MSA has five reporting groups, including Yukon River fall chum salmon. Genotype and analysis are further detailed by **XX et al**. We multiplied the mean annual estimated annual proportion of Yukon River fall chum salmon by the total index of juvenile Chum salmon to yield a fall Chum stock-specific juvenile index.

To translate the model estimated juvenile abundances to the scale of the observed juvenile abundances, we estimated a constant, *q*. We fit the model to the juvenile index ranging from brood years 2002 to 2022. Observation error was assumed as log-normally distributed:

Eq. 4.11

Total return, harvest and spawner estimates for Yukon River fall Chum salmon were provided by the Alaska Department of Fisheries and Game (ADFG) run reconstruction (Fleischman & Borba 2009). A full discussion of the run reconstruction model is available in Fleischman and Borba 2009. Model methods have remained the same, while the reconstruction model estimates have been updated with new data annually. Spawner abundances were estimated based on data from escapement monitoring projects that cover 95% of the drainage and yields drainage wide estimates. Harvest data to inform the run reconstruction was compiled from commercial fish tickets and subsistence surveys. In the reconstruction, the total run size was assumed as the sum of the spawners and harvest. We fit the model to the return, harvest and spawner estimates ranging from brood years 2002 to 2022. We assumed observation error was log-normally distributed for all stages. Here, observation error, , acts as a weighting factor for the likelihood, where smaller sigma’s contribute more to the complete likelihood which helps account for the contributions of different sources of information.

Eq. 4.12

Eq. 4.13

Eq. 4.14

Age composition data that informed the run reconstruction was collected from lower river fisheries and test fisheries (Fleischman & Borba 2009). We fit the estimation model to age composition estimates from each calendar year provided by the run reconstruction. We used predicted run size by calendar year and age (Nt+a+1,s=r,a) to calculate the predicted proportions at age by calendar year (). The difference between the annual return age composition predicted by the model, *,* and the observed return age composition, was minimized by relating the two through a multinomial distribution.

Eq. 4.15

#### 2.4 Priors

Priors were selected as generally weak or uninformative for a majority of parameters (Table 1). The exception were regularized priors applied to covariate coefficients, . We applied a prior with a normal distribution, mean of 0 and standard deviation of 0.1. This regularized prior was imposed to avoid spurious correlations among covariates and productivity estimates.

#### 2.5 Ecosystem Covariates

Covariates included in the IPM were gathered based on support from peer reviewed literature (Table 2, Figure 2). All covariates were mean-scaled and included in the IPM all at once.

##### 2.5.1 Egg to the first summer at sea

We considered four covariates hypothesized to impact juvenile salmon productivity from the egg stage to the end of their first summer at sea, including the mean spawner size trend from the parent generation, Yukon River flow rates, cumulative degree days for sea surface temperatures in the Northern Bering Sea and pollock recruitment index. We included the mean trend in spawner size at age for spawners that returned during the juvenile brood year. Size at age can impact reproduction potential and affect productivity. We hypothesized a positive relationship between size and productivity where bigger fish produce more offspring and have greater reproductive success (Ohlberger et al. 2020, Oke et al. 2020, Feddern et al. 2024). The Alaska Department of Fish and Game (ADFG) conducts standardized salmon escapement surveys across Alaska where they have recorded salmon length, sex and age since the 1990’s. This information is publicly available (Supplemental Table 1). We compiled Yukon River Chum salmon age and length data spanning 2000-2021 from project sites with current timeseries and estimated a common trend in size at age using a dynamic factor analysis (DFA) (Text S1.1, Table S**XX**).

We included the Yukon River mainstem mean flow for May and June. We hypothesized that increased river flow has a negative relationship with productivity as it makes juvenile foraging more difficult (Neuswanger et al. 2015). A majority of juvenile Chum leave the lower Yukon River Delta by the end of June and occasionally into July, depending on ice break up phenology (Miller & Weiss 2023). Given this outmigration timing, Yukon River flow rates in May and June are the most likely to impact juvenile feeding and address this hypothesis. We acquired monthly flow data (cubic feet per second) from a gage hosted by the USGS at Pilot Station, AK, along the Lower Yukon River (Table 2, Table S1).

We included Northern Bering Sea Summer (NBS) cumulative degree days (CDD) to represent the temperature conditions preceding the NBS survey and represent ecosystem conditions for the first couple months fish are at sea. We hypothesized a positive relationship between temperature and juvenile productivity, as suggested by empirical studies in the Bering Sea and bioenergetics modeling in Japan (Iino et al. 2022, Farley et al. 2024). The hypothesized positive relationship may arise if warmer temperatures enhance juvenile salmon growth rates as they enter the marine environment which can reduce size selective mortality and lead to greater survival (Beamish & Mahnken 2001, Farley et al. 2024). To calculate CDD we used the daily mean NBS SST, publicly available on the Alaska Fisheries Information Network (AKFIN). We summed temperature from June to August for each year brood year +1 to align with when juvenile salmonids would experience the temperature conditions.

We included the EBS walleye pollock (*Gadus chalcogrammus*) recruitment index from the pollock stock assessment to represent changes in salmon prey availability during the first summer at sea (Ianelli et al. 2023). While juvenile Chum salmon consume multiple fish species, forage fish typically have patchy distributions making them difficult to survey and estimate reliable indices, thus the pollock recruitment index represents robust prey index for juvenile Chum salmon, and typically are the most abundant species in the forage fish biomass (Hollowed et al. 2012). Young pollock represent a high-quality prey source for juvenile Chum (Farley & Moss 2009, Moss et al. 2009, Kaga et al. 2013). High quality prey sources are important for lipid accumulation and can lead to greater growth and productivity (Myers et al. 2009, Kaga et al. 2013, Farley et al. 2024).

##### 2.5.1 First winter at sea to terminal river harvest

We considered four covariates hypothesized to impact marine salmon productivity, which included the period from the end of the first summer at sea until the individuals are vulnerable to terminal harvest when they return to the Yukon River (Table 2, Figure 2). At the end of their first summer at sea, individuals leave the Bering Sea and typically migrate to the Gulf of Alaska and the Aleutian Islands, where they feed and mature. We included the following covariates when estimating survival for the marine stage, : fullness index, winter CDD from the Eastern Aleutian Islands, and annual total Chum and Pink salmon hatchery releases (separately) from Alaska, Japan, Korea and Russia. We included a juvenile stomach fullness index (SFI), to represent the conditions fish experience when they begin their first winter at sea, we hypothesized that a higher SFI, which represents better fish condition, would be positively related to adult productivity. The SFI is estimated from fullness data collected by the NBS survey (discussed in more detail below) and the fullness data collection methods are detailed in Murphy et al 2021. Stomach fullness data are collected from salmon at each station and recorded on a per station basis. To account for differences in the survey through space and time and account for differences in the number of stomachs examined at each station, we used a generalized additive model to estimate an SFI. The model took the following form:

Eq. 4.16

where is the expected log SFI, for the i-th observation in space and time. We included an intercept to estimate mean SFI, and a factor year effect, to estimate direct annual indices. is a spatial field represented by a tensor product of B-splines for geospatial coordinates (: latitude, : longitude), which allowed for anisotropy in the smoothing process. To generate standardized estimates used in the IPM, we predicted SFI across years for the mean survey latitude and longitude in the dataset (62.0°N, -168.1°W). The index model was assessed for convergence and residuals were assessed for homogeneity.

We included winter Eastern Aleutian CDD to represent the temperature conditions that young Yukon River Chum salmon experienced during their first winter at sea. This stage is hypothesized as a critical survival bottleneck in the lifecycle (Farley et al. 2024). To calculate CDD we used the daily mean E Aleutian SST, publicly available on the Alaska Fisheries Information Network (AKFiN), summed from November to February to represent winter conditions. We hypothesized a negative relationship between high CDD and productivity, as high temperatures can alter the prey base which is critical under higher metabolic demands of warm temperatures (Farley et al. 2024).

Finally, we included Chum and Pink hatchery release abundances, separately, from Alaska, Japan, Korea and Russia. We hypothesized a negative relationship between hatchery release abundances and adult marine productivity as increases in marine competition negatively impacts salmon stocks (Ruggerone et al. 2003, Cunningham et al. 2018, Scheuerell et al. 2020, Feddern et al. 2024). International hatchery release information is publicly available from the North Pacific Anadromous Fish Commission (Table S1). We use a lagged tolling average of hatchery releases so that they occur in the model at a timestep where releases would overlap with Fall chum salmon in the ocean. For example, Chum salmon hatchery releases in brood year 2000 could begin to interact with Yukon River Fall Chum salmon by 2002 (calendar year, t+2) and due to variation in age structure, brood year 2000 may continue to compete with wild Yukon River fall chum for 3 more years. To capture this, we include a three-year rolling average of hatchery releases from brood year t + 2. While migration distances and times certainly vary from these different release points, we include hatchery releases to represent a coarse marine competition index. Since Pink salmon have a different life history than Chum salmon, we did not use a rolling average, we added a t+1 lag so pink salmon from brood year 2000 could compete with Yukon River Fall Chum in calendar year 2001.

#### Supplemental Text:

##### Dynamic Factor Analysis for Spawner Size

To estimate a temporal trend for Yukon River Fall chum salmon mean size at age, we conducted a dynamic factor analysis (DFA) using data from the Yukon region collected between 2000-2021. Data from 2000 to 2016 was collated by Oke et al, and more recent years were downloaded from the Alaska Department of Fish and Game Age Sex Length Database (Alaska Department of Fish and Game 2024) for all escapement projects with recent data (Table S1). We filtered the data to include only escapement samples and marine ages ranging from 3-6 years. Prior to analysis, length measurements were standardized within each marine age group to have a mean of 0 and standard deviation of 1.

The DFA was implemented using the MARSS package in R to identify common trends in body size across different ocean residence times (Holmes et al. 2024). Our model treated the standardized lengths as linear combinations of latent trends and relates the observed data to overall estimated length trend. This was expressed in our observation model as:

where represents the observed data (with marine ages as rows and years as columns), **Z** is a matrix of factor loadings (0,1) on the latent trends that determines how observations relate to the estimated trends and represents observation errors. The observation errors were assumed to be multivariate normal with mean zero and variance-covariance matrix **R**.

The process model estimates the latent trend in spawner size across ages as a random walk with process error:

where is the state, or the spawner length trend, for each calendar year, y. represents process errors assumed to be multivariate normal with mean zero and variance-covariance matrix **Q**. Model parameters and states were estimated using maximum likelihood via the MARSS package, with model convergence set to a maximum of 1000 iterations.